1   **Metabolic adaptations in a range expanding arthropod**

2

3   Katrien H. P. Van Petegem [*a], David Renault [b], Robby Stoks [c] and Dries Bonte [a]

4

5   [a] Department of Biology, Ghent University, Ghent, Belgium

6   [b] UMR CNRS 6553 Ecobio, Université de Rennes 1, Rennes Cedex, France

7   [c] Evolution and Conservation, KU Leuven, Leuven, Belgium

8   [*] Corresponding author: Katrien.VanPetegem@ugent.be

9

10   **Running headline**: Metabolomics of a range expanding arthropod

11

12   **Summary**

13

14   Despite an increasing number of studies documenting life-history evolution during range

15   expansions or shifts, we lack a mechanistic understanding of the underlying physiological

16   processes. In this explorative study, we used a metabolomics approach to study

17   physiological changes associated with the recent range expansion of the two-spotted spider

18   mite (*Tetranychus urticae*). Mite populations were sampled along a latitudinal gradient from

19   range core to edge and reared under benign common garden conditions for two

20   generations. Using Gas Chromatography-Mass Spectrometry (GC-MS), we obtained

21   metabolic population profiles, which showed a gradual differentiation along the latitudinal

22   gradient, indicating (epi)genetic changes in the metabolome in association with range

23   expansion. These changes seemed not related with shifts in the mites' energetic

24   metabolism, but rather with differential use of amino acids. Particularly, more dispersive

25   northern populations showed lowered concentrations of several essential and non-essential

26   amino acids, suggesting a potential downregulation of metabolic pathways associated with

27   protein synthesis.

28

29   **Keywords**: essential amino acids, common garden, life-history evolution, GC-MS

30   metabolomics, global change, *Tetranychus urticae*

31

32   **Introduction**

33

34   During range expansions or range shifts, species' life histories can evolve on

35   ecological timescales (Phillips, Brown & Shine 2010). Changing environmental conditions

36   force species to locally adapt and spatial assortment of dispersive phenotypes leads to

37   increased dispersiveness at the expanding/shifting range edge (Shine, Brown & Phillips

38   2011). These evolutionary processes of local adaptation and spatial selection affect key life-

39   history traits like fecundity, development and dispersal (reviewed in Chuang & Peterson

40   2015). We therefore expect range edge populations to exhibit physiological adaptations that

41   underlie these observed trait evolutions. Such adaptations should be especially significant in

42   energy-producing pathways, and more particularly in glycolysis (Eanes 2011). Indeed, any

43   elevation of the performance of one life-history trait augments its energetic and metabolite

44   demands at the expense of other traits (*cfr.* the "Y" model of resource allocation, Van

45   Noordwijk & de Jong 1986; Zera & Harshman 2001), thus modifying the global metabolic

46   network operation. For instance, variations in lipid biosynthesis in association with dispersal

47   strategy highly impact metabolite fluxes through lipid pathways (see Zera 2011 for a review).

48   As a result, life-history differentiation is expected to be associated with changes in the

49   metabolome (*i.e.* the set of circulating metabolites within an organism, Oliver *et al.* 1998), as

50   was for example found for ageing in *Caenorhabditis elegans* (Fuchs *et al.* 2010) and

51   reproduction in the Malaria Mosquito *Anopheles gambiae* (Fuchs *et al.* 2014).

52

53   Metabolomics is a convenient technique which can be used as a candidate approach

54   to explore an organism's response to environmental variations and forthcoming

55   environmental changes (Hines *et al.* 2007; Miller 2007; Viant 2008; Bundy, Davey & Viant

56   2009; Lankadurai, Nagato & Simpson 2013; Hidalgo *et al.* 2014). It provides information on

57   the interaction between an organism's physiology and its natural environment by identifying

58   metabolites of low to moderate molecular mass within the whole body, cells, tissues or

59   biofluids. Compared to other –omics technologies like genomics and transcriptomics,

60   metabolomics has the significant advantage to focus on 'downstream' cellular functions

61   (Snart, Hardy & Barrett 2015), providing a more direct picture of the functional links

62   between causes and consequences of environmental variation (Foucreau *et al.* 2012).

63   Essentially, metabolomics can potentially provide a link between genotypes and phenotypes

64   (Fiehn 2002). When applied on individuals originating from different localities from range

65   core to edge, but reared for several generations under common garden conditions, it should

66   provide insights on the physiological adaptations that underlie life-history evolution during

67   range expansion.

68

69   Though a consideration of the whole-organism physiology allows a better

70   understanding of how life-history evolution in natural populations might occur and why this

71   evolution is sometimes constrained (Zera *et al.* 2001; 2011; Ricklefs & Wikelski 2002), few

72   studies documented metabolic variation in wild populations along natural gradients

73   (Sardans, Penuelas & Rivas-Ubach 2011). Instead, most studies assess plastic or evolutionary

74   responses to environmental stressors by manipulating abiotic variables in controlled

75   environments (Sardans *et al.* 2011; Colinet *et al.* 2012; Padfield *et al.* 2016). Notable

76   exceptions are the studies on *Arabidopsis lyrata* that demonstrated distinct metabolic

77   phenotypes along the species' latitudinal distribution, with a typical cold-induced

78   metabolome in the north, indicating adaptation to the local climate (Davey *et al.* 2008;

79   Davey, Woodward & Quick 2009), but no difference in metabolic fingerprint between large

80   connected *versus* marginal fragmented populations (Kunin *et al.* 2009). These studies,

81   however, used plants that were grown from seeds collected directly from the field.

82   Environmental maternal effects can therefore not be excluded.

83

84   The two-spotted spider mite, *Tetranychus urticae* Koch (Acari, Tetranychidae; Fig. 1),

85   a generalist pest species in greenhouses and orchards, expanded its European range from

86   the Mediterranean to at least southern Scandinavia (K. H. P. Van Petegem, personal

87   observation in 2011) during the last decades (for more information, see Carbonnelle *et al.*

88   2007). Previous research with *T. urticae* showed quantitative genetic life-history

89   differentiation along this latitudinal gradient, with daily fecundity, lifetime fecundity and

90   longevity decreasing from range core to edge, and egg survival, dispersal propensity and sex

91   ratio increasing from range core to edge (Van Petegem *et al.* 2016). We expected this life-

92   history differentiation to be associated with an evolution of the species' intermediary

93   metabolism, which would manifest into distinct metabolic phenotypes among populations

94   sampled along its expansion gradient.

95

96   Using a metabolomics approach, the current study aimed to test (i) whether the

97   metabolome of *T. urticae* evolved during its recent range expansion; *i.e.* whether a gradual

98   change in the mite's intermediary metabolism is present from range core to range edge,

99   showing in the appearance of progressively distinct metabolic phenotypes (metabotypes) (ii)

100   whether this metabolic differentiation could be associated with the up- or downregulation

101   of certain metabolic pathways, for instance enhanced glycolytic activities or lipid metabolism

102   and (iii) whether this evolutionary change in the species' metabotype is associated with the

103   life-history differentiation that has occurred during its range expansion.

104   **Materials and methods**

105

106   **Field sampling and common garden**

107   In August 2012, we hand-sampled mites from nine localities (one population per

108   locality) along an 800 km latitudinal gradient from northwestern Belgium to northern

109   Denmark (Fig. 2). Mites were found on infested leaves of *Lonicera periclymenum* (European

110   honeysuckle, five populations) at high latitudes and on *Euonymus europaeus* (European

111   spindle, one population), *Humulus lupulus* (common hop, one population), *Sambucus nigra*

112   (European black elderberry, one population) or *Lonicera periclymenum* (European

113   honeysuckle, one population) at lower latitudes. (More information is provided in appendix

114   A.1.) In the laboratory, fifty to several hundreds of mites per population were put on

115   separate whole bean plants (*Phaseolus vulgaris*, variety Prélude –a highly suitable host for *T.*

116   *urticae*, see Agrawal et al. 2002; Gotoh et al. 2004) and kept under controlled conditions at

117   room temperature with a light-regime of 16:8 LD. After one generation, ten adult female

118   mites per population were taken from their bean plant and put on a piece of bean leaf

119   ($\pm 30 cm^2$) on wet cotton in a Petri dish. Two such Petri dishes were prepared for each

120   population. The Petri dishes were then used to create a pool of synchronised two-day adult

121   female mites for each population (two-day adult females were preferred, since these are

122   significantly bigger than fresh adults). For this purpose, all females were allowed to lay eggs

123   during 24 hours in a climate room at 27 °C (an optimal temperature for our study species,

124   see Sabelis 1981), with a light-regime of 16:8 LD. The resulting same-aged eggs were

125   subsequently left to develop until they were two-day adult mites, of which only females

126   (which are easily visually recognised) were selected. As mites were kept in common garden

127   for two generations, all direct environmental effects were excluded. Furthermore, since the

128   common garden conditions were optimal for *T. urticae* (bean as host plant, 27°C, 16:8 LD

129   and relatively low mite densities), we could reasonably assume that potential differences

130   among metabotypes of mites from different origins (populations) did not result from

131   differential responses to restrictive/stressful rearing.

132

133   **Metabolomic profiling using Gas Chromatography-Mass Spectrometry (GC-MS)**

134   As we wanted to scan metabolites from different metabolite families (because of

135   their various but connected roles in general organismal physiology), we used GC-MS

136   metabolomics (Koek *et al.* 2011; Khodayari *et al.* 2013). For each population, we constructed

137   the metabolomic profile of five replicated pooled sets of fifty two-day-adult female mites.

138   Each set was placed in a microtube and directly transferred to -80 °C. To be able to measure

139   true quantities of metabolites, it is important to standardise the initial masses of each

140   extract. However, even when pooling fifty individuals, the masses of the replicates were too

141   low to be accurately measured (the measurement error for the mass of the microtube was

142   greater than the summed mass of the fifty mites). Yet, previous research showed that

143   female adult size does not differ among the nine sampled populations (Van Petegem *et al.*

144   2016). We could thus confidently use and interpret metabolite concentrations in

145   nmol/sample. The samples were first homogenised in ice-cold (-20 °C) methanol-chloroform

146   (2:1), using a tungsten-bead beating equipment (RetschTM MM301, Retsch GmbH, Haan,

147   Germany) at 25 Hz for 1.5 min. After addition of ice-cold ultrapure water, the samples were

148   centrifuged at 4,000 g for 5 min at 4 °C. The upper aqueous phase was then transferred to

149   new chromatographic glass vials, dried-out and resuspended in 30 µl of 20 mg L-1

150   methoxyamine hydrochloride (Sigma-Aldrich, St. Louis, MO, USA) in pyridine and incubated

151   under automatic orbital shaking at 40 °C for 60 min. Subsequently, 30 µl of N-methyl-N-

7

152 (trimethylsilyl) trifluoroacetamide (MSTFA; Sigma, #394866) was added and the

153 derivatisation was conducted at 40 °C for 60 min under agitation. The samples were then

154 analysed in a GC-MS system (Thermo Fischer Scientific Inc., Waltham, MA, USA), using the

155 same settings as in Khodayari *et al.* (2013). For this purpose, one microliter of each sample

156 was injected in the GC-MS system using the split mode (split ratio: 25:1). After that, the

157 selective ion monitoring (SIM) mode (electron energy: -70 eV) was used to search for the

158 sixty primary metabolites that are most often found in arthropod samples and that were

159 included in our spectral database (see appendix A.2 for a complete overview of these sixty

160 metabolites). The SIM mode ensured a precise annotation of the detected peaks. The

161 calibration curves were set using samples consisting of sixty pure reference compounds at

162 concentrations of 1, 2, 5, 10, 20, 50, 100, 200, 500, 750, 1000, 1500 and 2000 µM.

163 Chromatograms were deconvoluted using XCalibur v2.0.7 software (Thermo Fischer

164 Scientific Inc., Waltham, MA, USA). Finally, metabolite concentrations were quantified

165 according to their calibration curves.

166

167 **Statistics**

168 A total of forty-three metabolites were identified.

169 In a first step, we examined whether distinct metabotypes existed among the

170 sampled populations of *T. urticae* by running a MANOVA. Nine metabolites were first

171 removed from the dataset because they showed a more than 85% correlation with (an)other

172 metabolite(s). Then, all remaining metabolites were log-transformed to obtain a normal

173 distribution, which allowed to proceed to the MANOVA.

174 In a second step, we tested whether the metabolic profile was gradually changing as

175 a function of latitude (from range core to edge), or as a function of one of the six life-history

8

176   traits that were previously shown to covary with latitude (daily and lifetime fecundity, egg

177   survival, longevity, dispersal propensity and sex ratio; see Van Petegem *et al.* 2016, see also

178   appendix A.3). The concentrations of all forty-three metabolites were first auto-scaled and

179   transformed (the transformation that best fitted and normalised the data was retained: cube

180   root when looking for covariation with daily fecundity and egg survival; log for latitude; no

181   transformation for lifetime fecundity, longevity, dispersal propensity and sex ratio). Then,

182   metabolic differences among the nine populations were visualised using Partial Least

183   Squares – Discriminant Analysis (PLS-DA). This multivariate analysis was performed using

184   MetaboAnalyst 3.0 (Xia *et al.* 2009; Xia *et al.* 2012; Xia *et al.* 2015). By ordering the

185   populations according to their latitude or according to one of the six life-history traits

186   covarying with latitude, it was possible to check for trends in the metabolite concentrations.

187   To validate the significance of this interpopulation variation, permutation tests (2000

188   permutations) were run for replicates (with 5 replicates per population) using separation

189   distance (B/W) test statistics. The PLS-DA provided Variable Importance in Projection (VIP)

190   scores, which gave a first overview of the possible existence of a general pattern in the

191   concentrations of quantified metabolites along our invasion gradient: low VIP-scores depict

192   a weak and high scores a strong global pattern. Using a step-wise procedure, only those

193   metabolites with a VIP score of at least 1.2 (1.0 for egg survival because removing the

194   metabolites with a score between 1.0 and 1.2 resulted in a decreased percentage of

195   variation explained) for the first and/or second component were retained for further

196   analysis (compared to 0.8 in Tenenhaus 1998).

197   In a third step, univariate analyses were performed to test, metabolite by metabolite,

198   whether the global patterns obtained in the previous step could be confirmed. As we aimed

199   at determining if the metabolite levels showed latitudinal patterns along the expansion, we

9

200 did not run an ANOVA on individual metabolites, but rather processed regressions. Using

201 SAS 9.4 (SAS Institute Inc. 2013), the linear regressions were run for all influential

202 metabolites (VIP scores >1.2, except for egg survival, as mentioned above). Because all five

203 collected replicates per population originated from only one field sample, we ran the

204 regressions using population averages. As our study is explorative, we wanted to avoid false

205 negatives (with the chance of making a Type II error). We therefore did not correct for

206 multiple comparisons (*e.g.* Bonferroni correction). Given the large number of statistical tests,

207 the limited number of populations, and the fact that we highly smoothed differences among

208 populations by rearing the specimens under common garden conditions for two generations

209 (which is atypical for metabolomics studies, where organisms are usually subject to a

210 stressor to elicit a strong response), such a correction would have greatly diminished the

211 statistical power.

212 A final step linked the selected individual metabolites with one or more metabolic

213 pathways, thus identifying those pathways that were potentially up- or downregulated

214 during the range expansion of *T. urticae*. Pathway enrichment analyses were performed in

215 Metaboanalyst 3.0 (Xia *et al.* 2009; Xia *et al.* 2012; Xia *et al.* 2015) with those metabolites

216 that showed significant effects in the univariate analyses of step three. These pathway

217 analyses were performed with a Fisher's exact test algorithm, which we ran using the

218 metabolic pathways of *Drosophila melanogaster* (no closer relative of *T. urticae* was

219 available in the program, but primary metabolites are anyway highly conserved, especially

220 among non-blood feeding arthropods). The algorithm calculates the match (number of hits)

221 between the metabolites in a dataset and the totality of metabolites present in a specific

222 pathway. Furthermore, it uses a pathway topology analysis to compute a value for the

223   impact of these metabolites on the pathway. As multiple comparisons are made, corrected

224   Holm p-values are provided.

225

226   **Results**

227

228   *MANOVA*

229   The metabotypes of the female mites significantly differed among the nine sampled

230   populations ($F_{184,160}=2.2$, p<0.001). In further analyses, the potential covariation of

231   metabolite levels with latitude or one of the life-history traits covarying with latitude was

232   then assessed.

233

234   *Latitudinal covariation*

235   The PLS-DA showed a separation among the nine populations, which was visible on 3-

236   D score plots (see appendix A.4). Of the forty-three identified metabolites, seventeen had

237   VIP scores of at least 1.2 and were thus retained for further analysis (Fig. 3A). They showed a

238   clear general trend from high values in southern to low values in northern populations (Fig.

239   3A).

240   In the subsequent linear regressions, eleven of these seventeen metabolite

241   concentrations varied significantly: ten decreased and one increased with increasing latitude

242   (Fig. 3A and appendix A.6). Among these eleven metabolites, five essential amino acids,

243   three non-essential amino acids (see Rodriguez & Hampton (1966) for an overview of all

244   essential amino acids in *T. urticae* —we defined tryptophan, which is not included in this

245   overview, as essential) and one intermediate of the citric acid cycle can be mentioned.

246   Pathway analysis indicated that of these eleven metabolites, eight play a significant

247   role in the aminoacyl-tRNA biosynthesis (total: 67, hits: 8, impact=0, Holm p=2.7062E-6) and

12

248   four in the valine, leucine and isoleucine biosynthesis (total: 13, hits: 4, impact=0.9999, Holm

249   p=4.0129E-4) (both pathway maps are provided in appendix A.7).

250

251   *Life history covariation*

252   The PLS-DA showed a separation between the nine (eight for egg survival, for which

253   no data were available for population SVI) populations, which was visible on 3-D score plots

254   (see appendix A.4). Of the forty-three identified metabolites, only those which explained

255   most of the interpopulation variation for a certain life-history trait (high VIP score) were

256   retained for further analysis. Fourteen were retained for daily and sixteen for lifetime

257   fecundity, twenty were retained for egg survival, eleven for longevity, thirteen for dispersal

258   propensity and ten for sex ratio (Fig. 3B-G). Figure 3 (B-G) shows clear indications of a

259   positive correlation between lifetime fecundity and its sixteen selected metabolites. In

260   contrast, figure 3 (B-G) suggests a negative correlation between the twenty and thirteen

261   metabolites selected for, respectively, egg survival and dispersal propensity. For daily

262   fecundity, longevity and sex ratio, no clear trends were visible.

263   In the subsequent linear regressions, one significant correlation was found for egg

264   survival (a sugar alcohol) and seven for dispersal propensity (including four essential amino

265   acids and one sugar). No significant results were found for lifetime fecundity, daily fecundity,

266   longevity and sex ratio (Fig. 3B-G and appendix A.6).

267   Pathway analysis indicated that four of the seven metabolites which negatively

268   correlated with dispersal propensity play a significant role in the aminoacyl-tRNA

269   biosynthesis (total: 67, hits: 4, impact=0, Holm p=0.0417) (the pathway map is provided in

270   appendix A.7). No associated pathways were found for egg survival.

271

272   **Discussion**

273

274   Of the forty-three metabolites identified in the GC-MS analysis, eighteen correlated

275   with latitude and/or one or more life-history traits. More specifically, eleven covaried

276   positively or negatively with latitude, seven showed a negative correlation with dispersal

277   propensity and one showed a negative correlation with egg survival. Of the eighteen

278   different metabolites, eleven amino acids could be shown to play an important role in the

279   aminoacyl-tRNA biosynthesis and four in the valine, leucine and isoleucine biosynthesis (see

280   pathway maps provided in appendix A.7).

281

282   Contrary to our hypothesis, our results indicate that the life-history evolution which

283   occurred during the recent range expansion of *T. urticae* (Van Petegem *et al.* 2016) was not

284   associated with shifts in the mites' energetic metabolism, but rather with shifts in its

285   anabolism. While our spectral database contained eleven sugars, only one sugar (fructose)

286   accounted for the separation among populations. This suggests that the genes involved in

287   encoding the mite's energetic metabolism (*i.e.* glycolysis, citric acid cycle, which typically

288   involve sugars) have not been significantly affected during the range expansion of *T. urticae*.

289   Instead, the observed differentiation in the mites' metabolome probably involved

290   evolutionary changes in the mites' anabolism, where amino acids play a central role in the

291   metabolic turnover of proteins. In more northern and more dispersive populations, the

292   aminoacyl-tRNA biosynthesis was downregulated. In this pathway, aminoacyl-tRNA is

293   formed by charging tRNA with an amino acid. The aminoacyl-tRNA then serves as a substrate

294   in protein synthesis or plays one of its many other roles in, for example, cell wall formation

295   or antibiotic biosynthesis (Raina & Ibba 2014). In accordance, the valine, leucine and

296   isoleucine biosynthesis, important for protein synthesis as well (Ahmed & Khan 2006;

297   Tamanna & Mahmood 2014), was downregulated in more northern populations.

298

299   The affected amino acids showed decreased concentrations toward higher latitudes

300   and showed a negative correlation with the dispersal propensity of *T. urticae*, which

301   increases towards the north. While, in general, amino acids are considered fundamental for

302   egg production and thus fecundity (Tulisalo 1971; O'Brien, Fogel & Boggs 2002; Mevi-Schutz

303   & Erhardt 2005; Fuchs *et al.* 2014, but see Heagle *et al.* 2002), not a single correlation was

304   found for fecundity, despite a clear positive trend in the PLS-DA. Of the eleven affected

305   amino acids, eight were essential and three non-essential. While the non-essential amino

306   acids could have been synthesised *de novo* from glucose, the essential amino acids could

307   only have been supplied through the mite's diet (Rodriguez *et al.* 1966). Though all mites

308   were kept in common garden, mites from northern, more dispersive populations were found

309   to contain lower essential amino acid concentrations. In line with the recent finding of

310   Fronhofer & Altermatt (2015) that a dispersal-foraging trade-off leads to a reduced

311   exploitation of resources at range margins, our results could indicate that northern, more

312   dispersive mites evolved lower essential amino acid concentrations because they consume

313   less of their food source. We should, however, keep in mind that metabolites were

314   measured only at one point in time and from whole organism samples. We are therefore

315   missing the temporal fluctuations of the metabolome over a day, and our data therefore

316   represent only a snapshot of the existing balance in terms of metabolite demand among

317   metabolic pathways.

318

319   An important challenge for metabolomics is understanding the relative contribution

320   of environmental and genetic factors in shaping an organism's metabolic phenotype (Bundy

321   *et al.* 2009). In the current study, mites were kept in common garden for two generations,

322   during which they were reared under optimal, non-stressful conditions. In contrast with

323   most metabolomic studies, mites were thus not subjected to a stressor to elicit a strong

324   environmentally-induced   plastic   stress-response.   Therefore,   any   metabolomic

325   differentiation was expected to be far less pronounced compared to stress-exposure studies.

326   As plastic, environmentally driven field-differences among populations were largely levelled

327   out through the two generations in common garden, only genetic factors were retained.

328   Long-lasting transgenerational plasticity can, however, not be fully excluded and we

329   therefore refer further to (epi)genetic factors (Verhoeven et al. 2016). Though (epi)genetic

330   factors are generally considered less determining than environmental factors (Robinson *et*

331   *al.* 2007; Frank, Noerenberg & Engel 2009; Matsuda *et al.* 2012), our results demonstrate a

332   clear (epi)genetic signal of metabolic differentiation along *T. urticae*'s invasion gradient. We

333   acknowledge, however, that we cannot exclude neutral processes, like for example serial

334   bottlenecks –including surfing mutations (Travis *et al.* 2007; Klopfstein *et al.* 2016) as

335   potential (co)sources of the found latitudinal metabolomic patterns. Furthermore, as the

336   host plant species on which the mites were sampled in the field covaried with latitude (with

337   *L. periclymenum* typically in the north), this could also have influenced our results. Yet,

338   regression slopes barely differed between models including all populations (hence all host

339   plant species) or models run for the subset of six populations collected on *L. periclymenum*

340   only (see appendix A.8). This indicates that the latitudinal signal was independent of host

341   plant identity.

342

343   This explorative study specifically examined whether range expansion might result in

344   evolutionary changes in an organism's metabolism. Despite non-stressful common garden

345   conditions, approximately forty per cent of the identified metabolites showed (epi)genetic

346   differentiation among populations. The more dispersive northern mites exhibited lower

347   concentrations of several essential and non-essential amino acids, suggesting a

348   downregulation of certain pathways linked to protein synthesis. Though effects were subtle

349   (but see earlier), our results clearly indicate that the metabolome of *T. urticae* underwent

350   (epi)genetic changes during the species' recent range expansion.

18

351   **Acknowledgements**

355

356   **Data accessibility**

357   - Sample locations and life-history trait values: included in appendix

358   - Results of GC-MS analysis: will be deposited in the Dryad repository (and added to the

359   reference list)

360   - Outcomes of statistical analyses: included in appendix

361

362   **Figure legends**

363

364   **Figure 1**

365   Picture showing two females and one egg of the two-spotted spider mites

366   (*Tetranychus urticae* Koch; Acari, Tetranychidae).

367

368   **Figure 2**

369   The map shows the nine field collection sites, which are situated in Belgium, The

370   Netherlands and Denmark. The graph shows the yearly number of frost days and the

371   average yearly temperature for each collection site along the latitudinal gradient. These

372   climatic data were obtained from FetchClimate (Microsoft Research, Cambridge) and were

373   averaged over a period of 35 years (1980 to 2015). Below the graph, arrows for each of six

374   life-history traits depict their trend along the latitudinal gradient (increase, decrease). (For

375   more detailed information, see appendix A.1 and Van Petegem *et al.* 2015.)

376

377   **Figure 3**

378   Variable importance plots resulting from the multivariate analyses (PLS-DA) on the

379   metabolomic data. These plots list those metabolites that, based on their VIP score,

380   contribute the most to explaining the variation among the nine populations in our dataset

381   (ODK, KVS, CAS, LAU, HED, BLA, TVE; SVI, SKA). The metabolites are ordered from high to low

382   VIP scores for component 1 (an overview off all scores for component 1 and 2 is provided in

383   appendix A.5). The colour codes indicate the relative concentration of a given metabolic

384   compound for each population (green=low concentration, to red= high concentration). The

385   populations themselves are ordered according to their latitude (A), or from low values at the

20

386   left to high values at the right for a given life-history trait (daily (B) or lifetime (C) fecundity,

387   egg survival (D), longevity (E), dispersal propensity (F) or sex ratio (G)). For example, in Fig.

388   3A, LAU is the population with the highest concentration of proline and ODK is the

389   southernmost population (lowest latitude). Below each column (population), a letter

390   signifies the host plant species from which mites were sampled in this population (*H.= H.*

391   *lupulus, S.= S. nigra, L.= L. periclymenum, E.= E. europaeus*). At the bottom of each plot, the

392   p-value resulting from the permutation test is given. At the left side of each plot, an asterisk

393   next to a metabolite name indicates a significant correlation between this metabolite and

394   latitude or the denoted life-history trait. For example, in Fig. 3A, proline shows a significant

395   negative correlation with latitude. (A detailed overview, including p- and F-values, of all

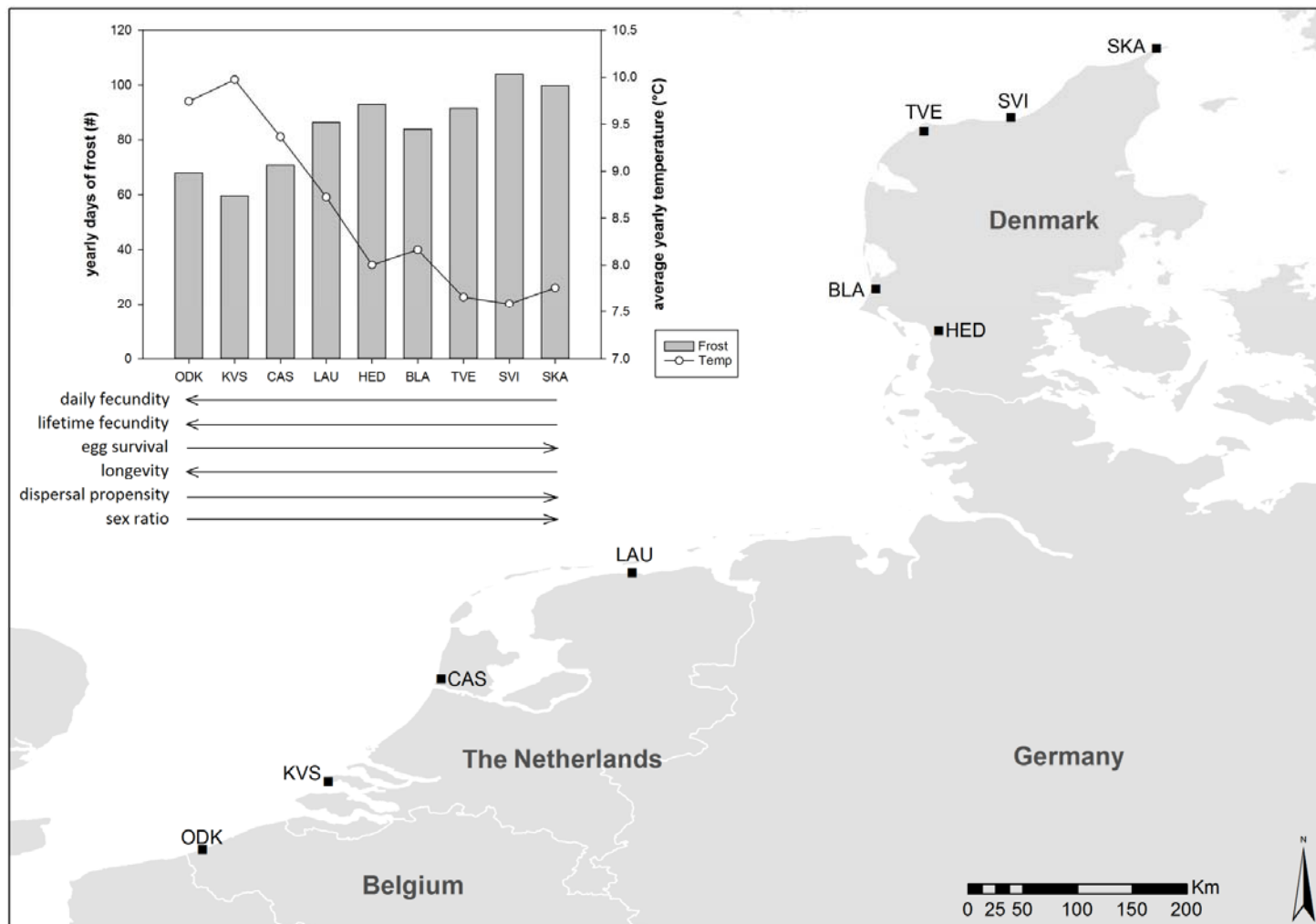396   linear regressions is found in appendix A.6.)

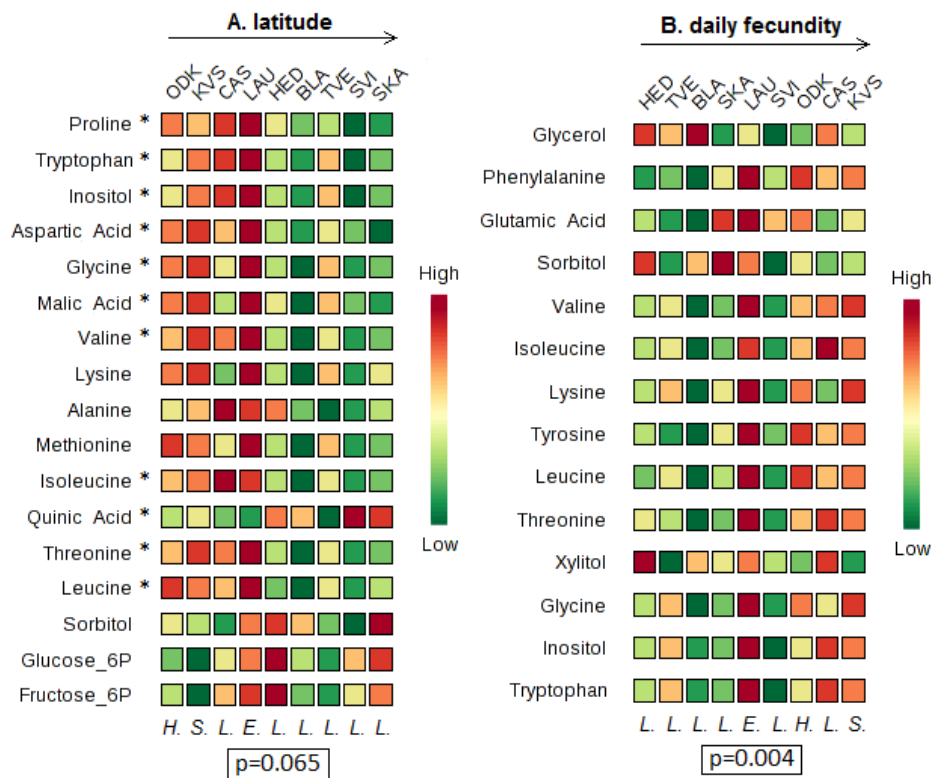397

398

399 **Figures**

400
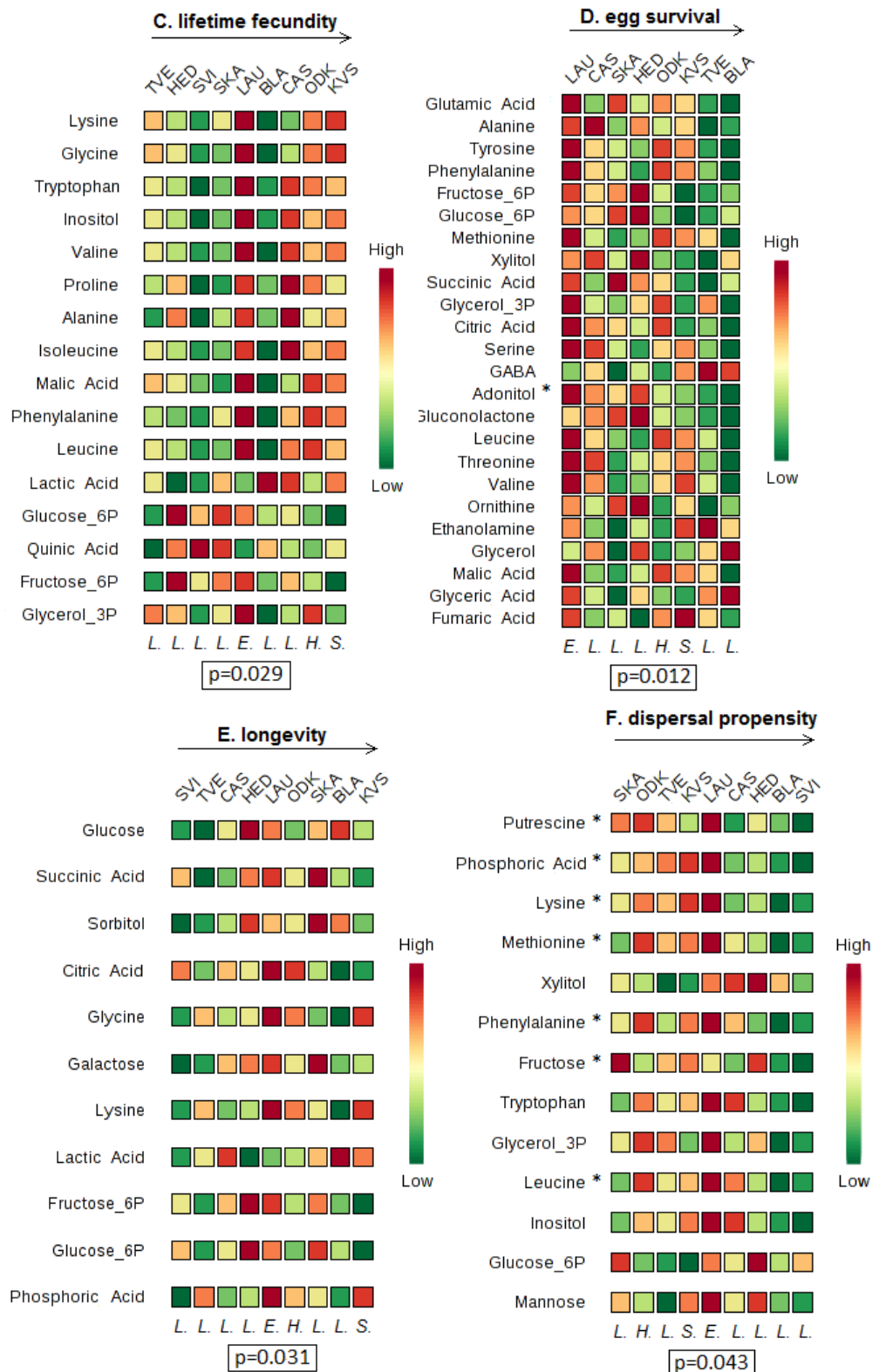
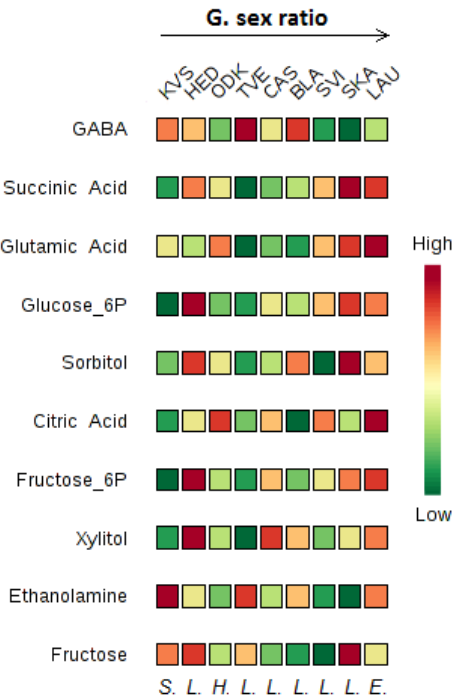401 **Figure 1**

402

403 **Figure 2**

404

405   **Figure 3**

406

407

408

409

410

411 **References**

412

413 Agrawal, A.A., Vala, F. & Sabelis, M.W. (2002). Induction of preference and performance

414   after acclimation to novel hosts in a phytophagous spider mite: Adaptive plasticity?

415   *American Naturalist,* **159 (5)**, 553-565.

416 Ahmed, I. & Khan, M.A. (2006) Dietary branched-chain amino acid valine, isoleucine and

417   leucine requirements of fingerling Indian major carp, *Cirrhinus mrigala* (Hamilton).

418   *British Journal of Nutrition,* **96 (3),** 450-460.

419 Bundy, J.G., Davey, M.P. & Viant, M.R. (2009) Environmental metabolomics: a critical review

420   and future perspectives. *Metabolomics,* **5 (1),** 3-21.

421 Carbonnelle, S., Hance, T., Migeon, A., Baret, P., Cros-Arteil, S. & Navajas, M. (2007)

422   Microsatellite markers reveal spatial genetic structure of *Tetranychus urticae* (Acari :

423   Tetranychidae) populations along a latitudinal gradient in Europe. *Experimental and*

424   *Applied Acarology,* **41 (4),** 225-241.

425 Chuang, A. & Peterson, C.R. (2016) Expanding population edges: theories, traits, and trade-

426   offs. *Global Change Biology,* **22 (2)**, 494-512.

427 Colinet, H., Larvor, V., Laparie, M. & Renault, D. (2012) Exploring the plastic response to cold

428   acclimation through metabolomics. *Functional Ecology,* **26 (3),** 711-722.

429 Davey, M.P., Burrell, M.M., Woodward, F.I. & Quick, W.P. (2008) Population-specific

430   metabolic phenotypes of *Arabidopsis lyrata* ssp. *petraea. New Phytologist,* **177 (2),**

431   380-388.

432 Davey, M.P., Woodward, F.I. & Quick, W.P. (2009) Intraspecfic variation in cold-temperature

433   metabolic phenotypes of *Arabidopsis lyrata* ssp. *petraea. Metabolomics,* **5 (1),** 138-

434   149.

27

435   Eanes, W.F. (2011) Molecular population genetics and selection in the glycolytic pathway.

436   *Journal of Experimental Biology,* **214 (2)**, 165-171.

437   Fiehn, O. (2002) Metabolomics - the link between genotypes and phenotypes. *Plant*

438   *Molecular Biology,* **48 (1-2),** 155-171.

439   Foucreau, N., Renault, D., Hidalgo, K., Lugan, R. & Petillon, J. (2012) Effects of diet and

440   salinity on the survival, egg laying and metabolic fingerprints of the ground-dwelling

441   spider *Arctosa fulvolineata* (Araneae, Lycosidae). *Comparative Biochemistry and*

442   *Physiology a-Molecular & Integrative Physiology,* **163 (3-4),** 388-395.

443   Frank, T., Noerenberg, S. & Engel, K.-H. (2009) Metabolite Profiling of Two Novel Low Phytic

444   Acid (lpa) Soybean Mutants. *Journal of Agricultural and Food Chemistry,* **57 (14),**

445   6408-6416.

446   Fronhofer, E.A. & Altermatt, F. (2015) Eco-evolutionary feedbacks during experimental range

447   expansions. *Nature Communications,* **6**.

448   Fuchs, S., Behrends, V., Bundy, J.G., Crisanti, A. & Nolan, T. (2014) Phenylalanine Metabolism

449   Regulates Reproduction and Parasite Melanization in the Malaria Mosquito. *Plos*

450   *One,* **9 (1)**.

451   Fuchs, S., Bundy, J.G., Davies, S.K., Viney, J.M., Swire, J.S. & Leroi, A.M. (2010) A metabolic

452   signature of long life in *Caenorhabditis elegans*. *Bmc Biology,* **8**.

453   Gotoh, T., Suwa, A., Kitashima, Y. & Rezk, H.A. (2004) Developmental and reproductive

454   performance of *Tetranychus pueraricola* Ehara and Gotoh (Acari : Tetranychidae) at

455   four constant temperatures. *Applied Entomology and Zoology,* **39 (4)**, 675-682.

456   Heagle, A.S., Burns, J.C., Fisher, D.S. & Miller, J.E. (2002) Effects of carbon dioxide

457   enrichment on leaf chemistry and reproduction by twospotted spider mites (Acari :

458   Tetranychidae) on white clover. *Environmental Entomology,* **31 (4),** 594-601.

459   Hidalgo, K., Mouline, K., Mamai, W., Foucreau, N., Dabire, K., Bouchereau, A., Simard, F. &

460   Renault, D. (2014) Novel insights into the metabolic and biochemical underpinnings

461   assisting dry-season survival in female malaria mosquitoes of the *Anopheles gambiae*

462   complex. *Journal of Insect Physiology,* **70,** 102-116.

463   Hines, A., Oladiran, G.S., Bignell, J.P., Stentiford, G.D. & Viant, M.R. (2007) Direct sampling of

464   organisms from the field and knowledge of their phenotype: Key recommendations

465   for environmental metabolomics. *Environmental Science & Technology,* **41 (9),** 3375-

466   3381.

467   Khodayari, S., Moharramipour, S., Larvor, V., Hidalgo, K. & Renault, D. (2013) Deciphering the

468   Metabolic Changes Associated with Diapause Syndrome and Cold Acclimation in the

469   Two-Spotted Spider Mite *Tetranychus urticae*. *Plos One,* **8 (1).**

470   Klopfstein, S., Currat, M., Excoffier, L. (2006) The fate of mutations surfing on the wave of a

471   range expansion. *Molecular Biology and Evolution,* **23 (3),** 482-490.

472   Koek, M.M., Jellema, R.H., van der Greef, J., Tas, A.C. & Hankemeier, T. (2011) Quantitative

473   metabolomics   based   on   gas   chromatography   mass   spectrometry: status   and

474   perspectives. *Metabolomics,* **7 (3),** 307-328.

475   Kunin, W.E., Vergeer, P., Kenta, T., Davey, M.P., Burke, T., Woodward, F.I., Quick, P.,

476   Mannarelli, M.-E., Watson-Haigh, N.S. & Butlin, R. (2009) Variation at range margins

477   across multiple spatial scales: environmental temperature, population genetics and

478   metabolomic phenotype. *Proceedings of the Royal Society B-Biological Sciences,* **276**

479   **(1661),** 1495-1506.

480   Lankadurai, B.P., Nagato, E.G. & Simpson, M.J. (2013) Environmental metabolomics: an

481   emerging   approach   to   study   organism   responses   to   environmental   stressors.

482   *Environmental Reviews,* **21 (3),** 180-205.

483   Matsuda, F., Okazaki, Y., Oikawa, A., Kusano, M., Nakabayashi, R., Kikuchi, J., Yonemaru, J.-I.,

484   Ebana, K., Yano, M. & Saito, K. (2012) Dissection of genotype-phenotype associations

485   in rice grains using metabolome quantitative trait loci analysis. *Plant Journal,* **70 (4),**

486   624-636.

487   Mevi-Schutz, J. & Erhardt, A. (2005) Amino acids in nectar enhance butterfly fecundity: A

488   long-awaited link. *American Naturalist,* **165 (4),** 411-419.

489   Miller, M.G. (2007) Environmental metabolomics: A SWOT analysis (strengths, weaknesses,

490   opportunities, and threats). *Journal of Proteome Research,* **6 (2),** 540-545.

491   O'Brien, D.M., Fogel, M.L. & Boggs, C.L. (2002) Renewable and nonrenewable resources:

492   Amino acid turnover and allocation to reproduction in Lepidoptera. *Proceedings of*

493   *the National Academy of Sciences of the United States of America,* **99 (7),** 4413-4418.

494   Oliver, S.G., Winson, M.K., Kell, D.B. & Baganz, F. (1998) Systematic functional analysis of the

495   yeast genome. *Trends in Biotechnology,* **16 (9),** 373-378.

496   Padfield, D., Yvon-Durocher, G., Buckling, A., Jennings, S. & Yvon-Durocher, G. (2016) Rapid

497   evolution of metabolic traits explains thermal adaptation in phytoplankton. *Ecology*

498   *Letters,* **19 (2),** 133-142.

499   Phillips, B.L., Brown, G.P. & Shine, R. (2010) Life-history evolution in range-shifting

500   populations. *Ecology,* **91 (6),** 1617-1627.

501   Raina, M. & Ibba, M. (2014) tRNAs as regulators of biological processes. *Frontiers in Genetics,*

502   **5,** 171.

503   Ricklefs, R.E. & Wikelski, M. (2002) The physiology/life-history nexus. *Trends in Ecology &*

504   *Evolution,* **17 (10),** 462-468.

505 Robinson, A.R., Ukrainetz, N.K., Kang, K.-Y. & Mansfield, S.D. (2007) Metabolite profiling of

506 Douglas-fir (*Pseudotsuga menziesii*) field trials reveals strong environmental and

507 weak genetic variation. *New Phytologist,* **174 (4),** 762-773.

508 Rodriguez, J.G. & Hampton, R.E. (1966) Essential amino acids determined in 2-spotted spider

509 mite *Tetranychus urticae* KOCH (Acarina, Tetranychidae) with glucose-U-C[14]. *Journal*

510 *of Insect Physiology,* **12 (10),** 1209-1216.

511 Sabelis, M.W. (1981) Biological control of two-spotted spider mites using phytoseiid

512 predators. Part I. Modelling the predator-prey interaction at the individual level.

513 Department of Theoretical Production Ecology, Department of Entomology.

514 Wageningen, Agricultural University. Doctor in Agricultural Sciences: 242.

515 Sardans, J., Penuelas, J. & Rivas-Ubach, A. (2011) Ecological metabolomics: overview of

516 current developments and future challenges. *Chemoecology,* **21 (4),** 191-225.

517 Shine, R., Brown, G.P. & Phillips, B.L. (2011) An evolutionary process that assembles

518 phenotypes through space rather than through time. *Proceedings of the National*

519 *Academy of Sciences of the United States of America,* **108 (14),** 5708-5711.

520 Snart, C.J.P., Hardy, I.C.W. & Barrett, D.A. (2015) Entometabolomics: applications of modern

521 analytical techniques to insect studies. *Entomologia Experimentalis Et Applicata,* **155**

522 **(1),** 1-17.

523 Tamanna, N. & Mahmood, N. (2014) Emerging roles of branched-chain amino acid

524 supplementation in human diseases. *International Scholarly Research Notices,* **2014**.

525 Tenenhaus, M. (1998) *La régression PLS*. Editions Technip, Paris.

526 Travis, M.J., Muenkemueller, T., Burton, O.J., Best, A., Dytham, C., Johst, K. (2007)

527 Deleterious mutations can surf to high densities on the wave front af an expanding

528 population. *Molecular Biology and Evolution,* **24 (10)**, 2334-2343.

529   Tulisalo, U. (1971) Free and bound amino acids of three host plant species and various

530   fertilizer treatments affecting the fecundity of the two-spotted spider mite,

531   *Tetranychus urticae* Koch (Acarina, Tetranychidae). *Annales Entomologicae Fennicae,*

532   **37,** 155-163.

533   Van Noordwijk, A.J. & de Jong, G. (1986) Acquisition and allocation of resources - Their

534   influence on variation in life-history tactics. *American Naturalist,* **128 (1),** 137-142.

535   Van Petegem, K.H.P., Boeye, J., Stoks, R. & Bonte, D. (2016) Spatial selection and local

536   adaptation jointly shape life-history evolution during range expansion. bioRxiv

537   031922; doi: http://dx.doi.org/10.1101/031922.

538   Verhoeven, K.J.F., vonHoldt, B.M. & Sork, V.L. (2016) Epigenetics in ecology and evolution:

539   what we know and what we need to know. *Molecular Ecology,* **25 (8)**, 1631-1638.

540   Viant, M.R. (2008) Recent developments in environmental metabolomics. *Molecular*

541   *Biosystems,* **4 (10),** 980-986.

542   Xia, J., Mandal, R., Sinelnikov, I.V., Broadhurst, D. & Wishart, D.S. (2012) MetaboAnalyst

543   2.0—a comprehensive server for metabolomic data analysis. *Nucleic Acids Research.*

544   **40 (W1)**, W127-W133.

545   Xia, J., Psychogios, N., Young, N. & Wishart, D.S. (2009) MetaboAnalyst: a web server for

546   metabolomic data analysis and interpretation. *Nucleic Acids Research,* **37 (suppl2),**

547   W652-W660.

548   Xia, J., Sinelnikov, I.V., Han, B. & Wishart, D.S. (2015) MetaboAnalyst 3.0—making

549   metabolomics more meaningful. *Nucleic Acids Research.* **43 (W1)**, W251-W257.

550   Zera, A.J. (2011) Microevolution of intermediary metabolism: evolutionary genetics meets

551   metabolic biochemistry. *Journal of Experimental Biology,* **214 (2)**, 179-190.

552   Zera, A.J. & Harshman, L.G. (2001) The physiology of life history trade-offs in animals. *Annual*

553   *Review of Ecology and Systematics,* **32,** 95-126.

554

555

556 **Appendix**

557

558 **A.1: Overview of the field collection sites and their respective life-history trait values**

559 An overview of the localities where *T. urticae* was sampled in the field, together with population-level life-history trait values (daily fecundity (#eggs/day), lifetime fecundity (total #eggs), egg survival (%),

560 longevity (#days), dispersal propensity (%) and sex ratio (#male/total)). These values, originating from a previous study (Van Petegem *et al.* 2016), were used as the independent values in our linear regressions.

561 Field collection sites were located along a latitudinal gradient, spanning the coasts of Belgium (BEL), The Netherlands (NTL), Germany and Denmark (DEN). The denoted plant species (*Lonicera periclymenum*,

562 *Euonymus europaeus*, *Sambucus nigra* and *Humulus lupulus*) are the species on which the mites were found and sampled.

code	collection site	city, country	Coordinates (WGS 84)	plant species	daily fecundity	lifetime fecundity	egg survival	longevity	dispersal propensity	sex ratio
SKA	Flagbakkevej	Skagen, DEN	57.72, 10.53	*L. periclymenum*	3.85	33.42	96.72	10.80	9.47	0.33
SVI	Sletteåvej	Fjerritslev, DEN	57.15, 9.33	*L. periclymenum*	5.19	29.33	.	6.33	41.41	0.32
TVE	Hindingvej	Thisted, DEN	57.04, 8.62	*L. periclymenum*	3.65	23.42	100.00	7.42	14.53	0.26
BLA	V. Hennebysvej	Henne, DEN	55.74, 8.22	*L. periclymenum*	3.72	40.92	100.00	11.00	39.47	0.29
HED	picnic-place along 11/24	Gredstedbro, DEN	55.39, 8.74	*L. periclymenum*	3.48	29.17	96.88	9.60	31.45	0.22
LAU	Schildhoek	Lauwersoog, NTL	53.41, 6.22	*E. europaeus*	4.25	37.00	87.30	9.75	16.75	0.34
CAS	Hoofdweg	Heemskerk, NTL	52.53, 4.65	*L. periclymenum*	5.54	41.75	96.15	9.40	19.21	0.28
KVS	Bosweg	Burgh-Haamstede, NTL	51.68, 3.72	*S. nigra*	6.34	98.67	99.02	17.55	15.03	0.19
ODK	Duinparklaan	Koksijde, BEL	51.12, 2.68	*H. lupulus*	5.22	50.42	98.99	10.08	9.81	0.25

563

564 **A.2: Overview of the 60 metabolites included in our spectral database**

565 A list of all 60 metabolites screened for in our single quadrupole GC-MS. Each of these metabolites can be characterised by several ions

566 of different masses. We used the selective ion monitoring mode (SIM, see Waller *et al.* 2007) instead of the full scan mode. In this SIM

567 mode, the MS instrument is set to only look for specific previously-known ion masses of interest rather than to screen for all masses

568 over a wide range. The instrument can therefore be very specific for a particular compound of interest. Our original database contained

569 spectral information for 85 primary metabolites from a range of plant and invertebrate models. After removing all metabolites specific

570 to plant models and those (polyamines for instance) having concentrations lower than the detection limit of our equipment, only the 60

571 specific metabolites listed in the table below remained in our spectral database. Metabolites are classified according to six categories:

572 amino acid, polyol, sugar, intermediate of the citric acid cycle, 'other' metabolite and metabolite not found in any of our samples.

category	metabolite
amino acid	Alanine
amino acid	Aspartic acid
amino acid	Glutamic acid
amino acid	Glycine
amino acid	Isoleucine
amino acid	Leucine
amino acid	Lysine
amino acid	Methionine
amino acid	Phenylalanine
amino acid	Proline
amino acid	Serine
amino acid	Threonine
amino acid	Tryptophan
amino acid	Tyrosine
amino acid	Valine
intermediate citric acid cycle	Citric acid
intermediate citric acid cycle	Fumaric acid
intermediate citric acid cycle	Malic acid
intermediate citric acid cycle	Succinic acid
polyol	Adonitol
polyol	Arabitol
polyol	Glycerol
polyol	Glycerol-3-phosphate
polyol	Inositol
polyol	Sorbitol
polyol	Xylitol
sugar	Fructose
sugar	Fructose-6-phosphate
sugar	Galactose
sugar	Glucose

| sugar	Glucose-6-phosphate
sugar	Mannose
sugar	Ribose
other	Ethanolamine
other	GABA
other	Gluconolactone
other	Glyceric acid
other	Lactic acid
other	Ornithine
other	Phosphoric acid
other	Putrescine
other	Quinic acid
other	Spermidine
not found	Arabinose
not found	Asparagine
not found	Cadaverine
not found	Citrulline
not found	Cytosine
not found	Erythritol
not found	Galactitol
not found	Galacturonic acid
not found	Maltose
not found	Mannitol
not found	Octopamine
not found	Pipecolic acid
not found	Saccharose
not found	Spermine
not found	Trehalose
not found	Tyramine
not found	Xylose

575 **A.3: correlation matrix**

576 LAT (latitude), LIFE (lifetime fecundity), DAFE (daily fecundity), DISP (dispersal propensity), EGSU (egg survival), LONG (longevity), DIAP

577 (diapause incidence) and SERA (sex ratio). All significant correlations (Bonferroni correction not implemented) are in bold. As only a

578 selection of populations was included in the current study, the strength of the correlations between latitude and the shown life-history

579 traits decreased relative to the correlations found in Van Petegem *et al.* (2016).

580

Pearson Correlation Coefficients							
Prob > |r| under H0: Rho=0							
Number of Observations							
	LAT	LIFE	DAFE	DISP	EGSU	LONG	DIAP	SERA
LAT	1	-0.67384	-0.67452	0.38101	0.13882	-0.52688	0.47464	0.43955
		0.0466	**0.0463**	0.3117	0.743	0.145	0.1967	0.2365
	9	9	9	9	8	9	9	9
LIFE	-0.67384	1	0.75065	-0.27469	0.18641	0.9203	-0.337	-0.5922
	0.0466		**0.0198**	0.4744	0.6585	**0.0004**	0.3752	0.0929
	9	9	9	9	8	9	9	9
DAFE	-0.67452	0,75065	1	-0.18302	0.06919	0.46919	0.13277	-0.33554
	0.0463	**0.0198**		0.6374	0.8707	0.2026	0.7335	0.3774
	9	9	9	9	8	9	9	9
DISP	0.38101	-0.27469	-0.18302	1	0.14537	-0.31841	0.66023	0.11912
	0.3117	0.4744	0.6374		0.7313	0.4037	0.0529	0.7602
	9	9	9	9	8	9	9	9
EGSU	0.13882	0.18641	0.06919	0.14537	1	0.16269	0.19129	-0.59346
	0.743	0.6585	0.8707	0.7313		0.7003	0.65	0.1209
	8	8	8	8	8	8	8	8
LONG	-0.52688	0.9203	0.46919	-0.31841	0.16269	1	-0.52876	-0.54739
	0.145	**0.0004**	0.2026	0.4037	0.7003		0.1433	0.1271
	9	9	9	9	8	9	9	9
DIAP	0.47464	-0.337	0.13277	0.66023	0.19129	-0.52876	1	0.29697
	0.1967	0.3752	0.7335	0.0529	0.65	0.1433		0.4377
	9	9	9	9	8	9	9	9
SERA	0.43955	-0.5922	-0.33554	0.11912	-0.59346	-0.54739	0.29697	1
	0.2365	0.0929	0.3774	0.7602	0.1209	0.1271	0.4377	
	9	9	9	9	8	9	9	9

581

582

583   **A.4: 3-D score plots**

584   3-D score plots resulting from the multivariate analyses (PLS-DA) performed on our metabolomic data are given for latitude (A) and each

585   of six life-history traits (daily fecundity (B), lifetime fecundity (C), egg survival (D), longevity (E), dispersal propensity (F) and sex ratio

586   (G)). The mean metabolic phenotypes for each replicate (five in total) of all nine populations (eight for egg survival) are represented by a

587   dot and arranged in space according to their projections on three component axes. Each of these axes explains a certain amount (%) of

588   the metabolic variation that is present in the dataset. More similar points (populations with a similar metabolome) are placed closer

589   together.

590

591

C. lifetime fecundity

592

593

594

F. dispersal propensity

595

596

597
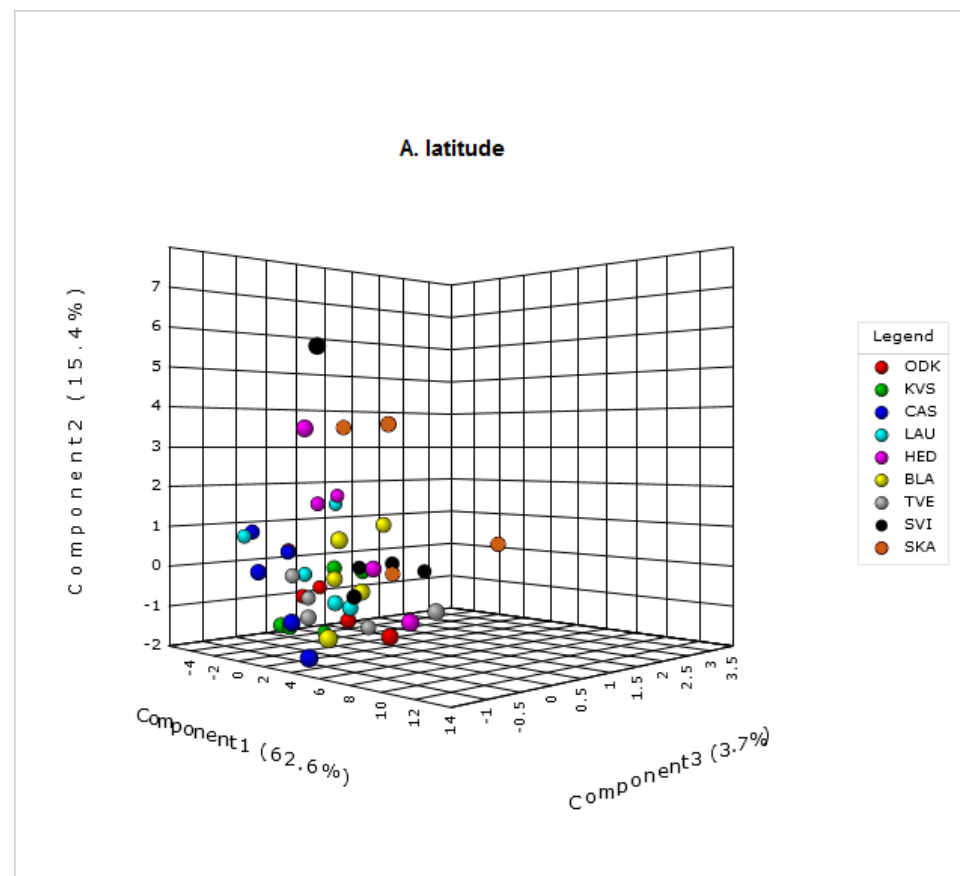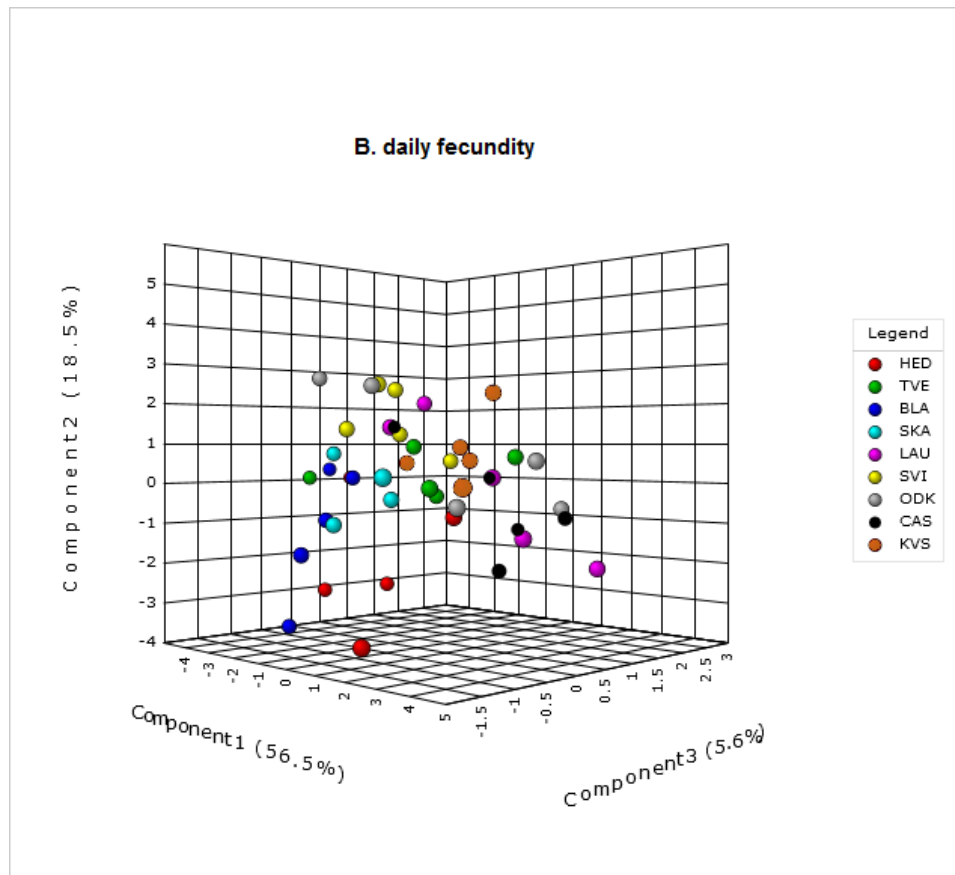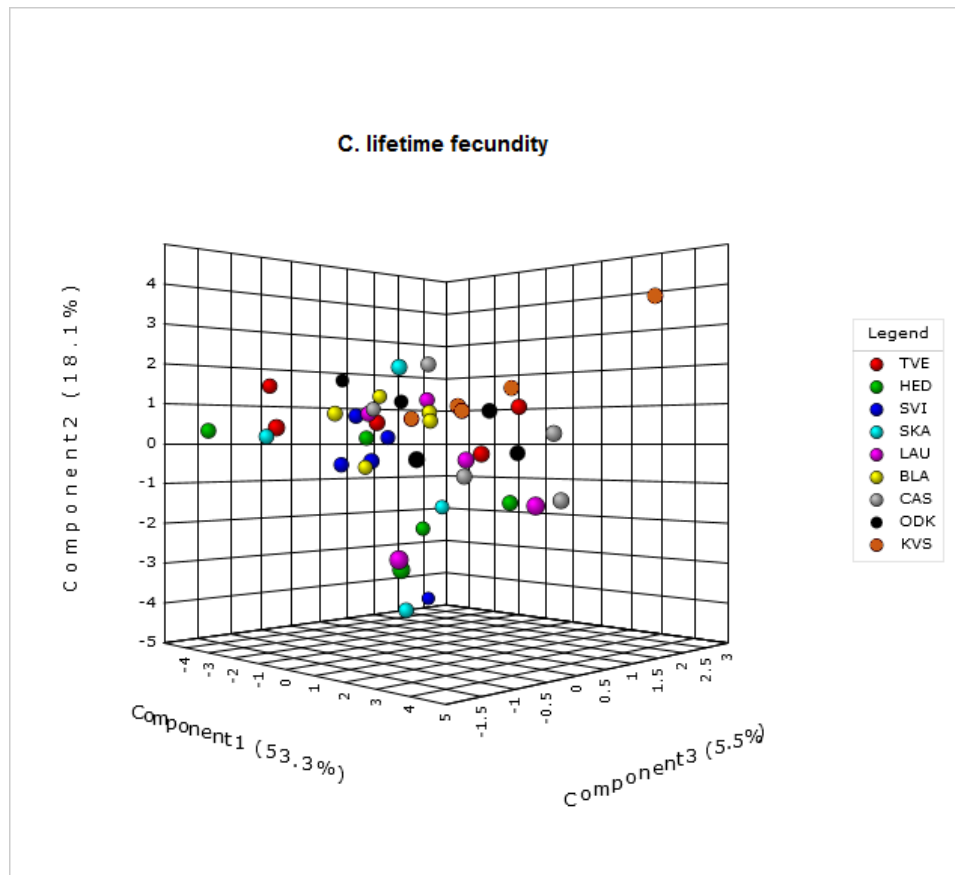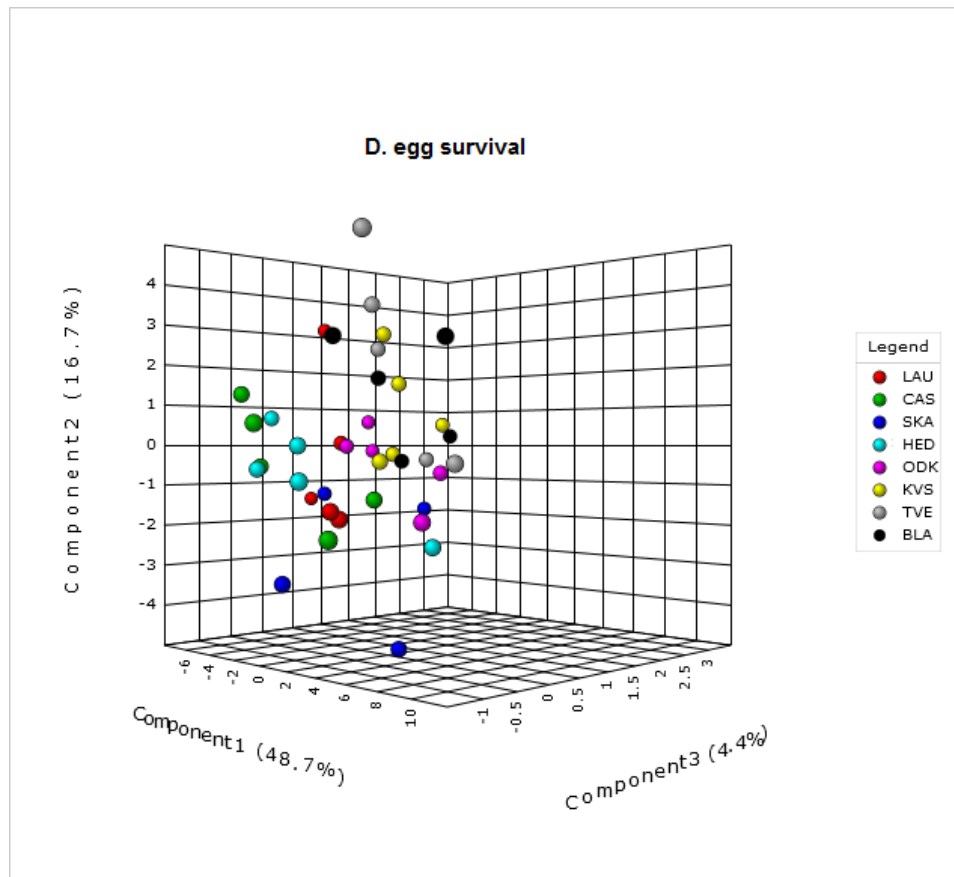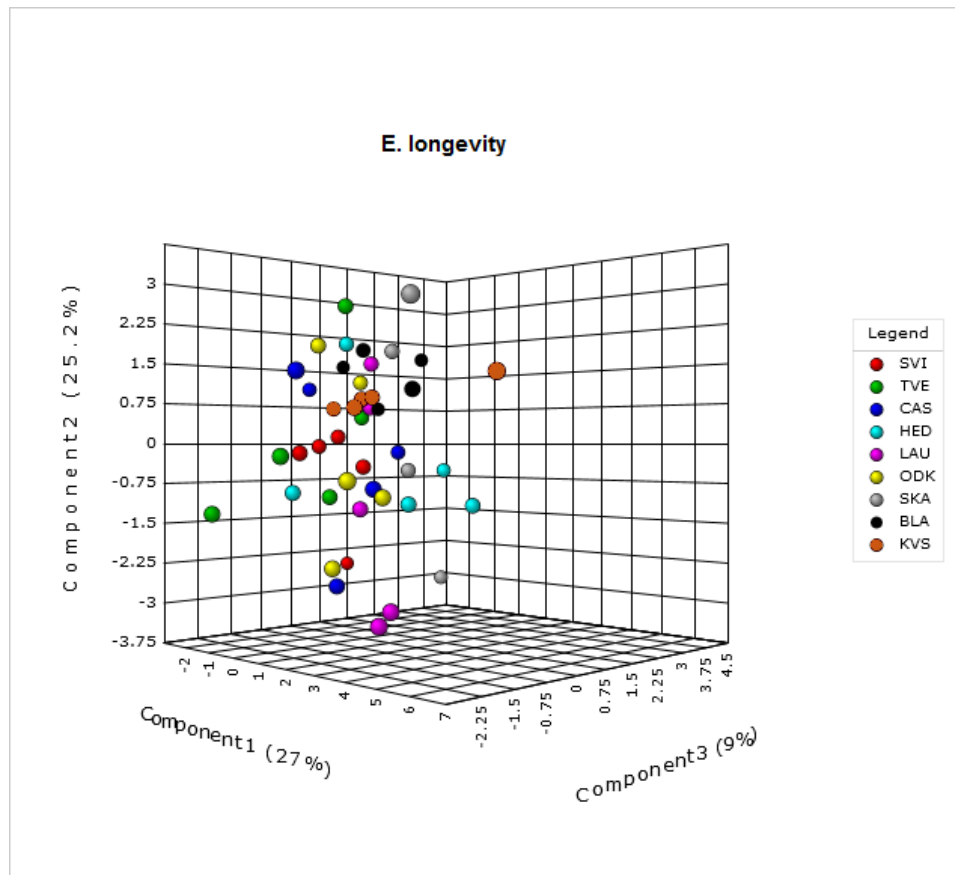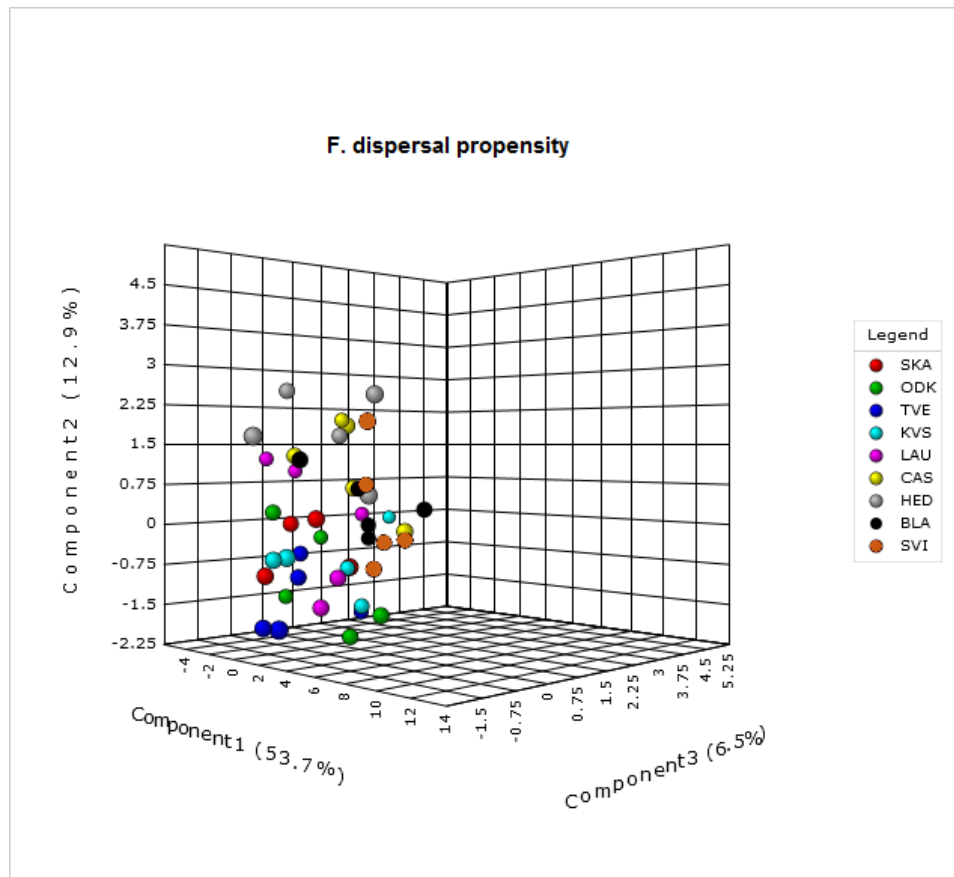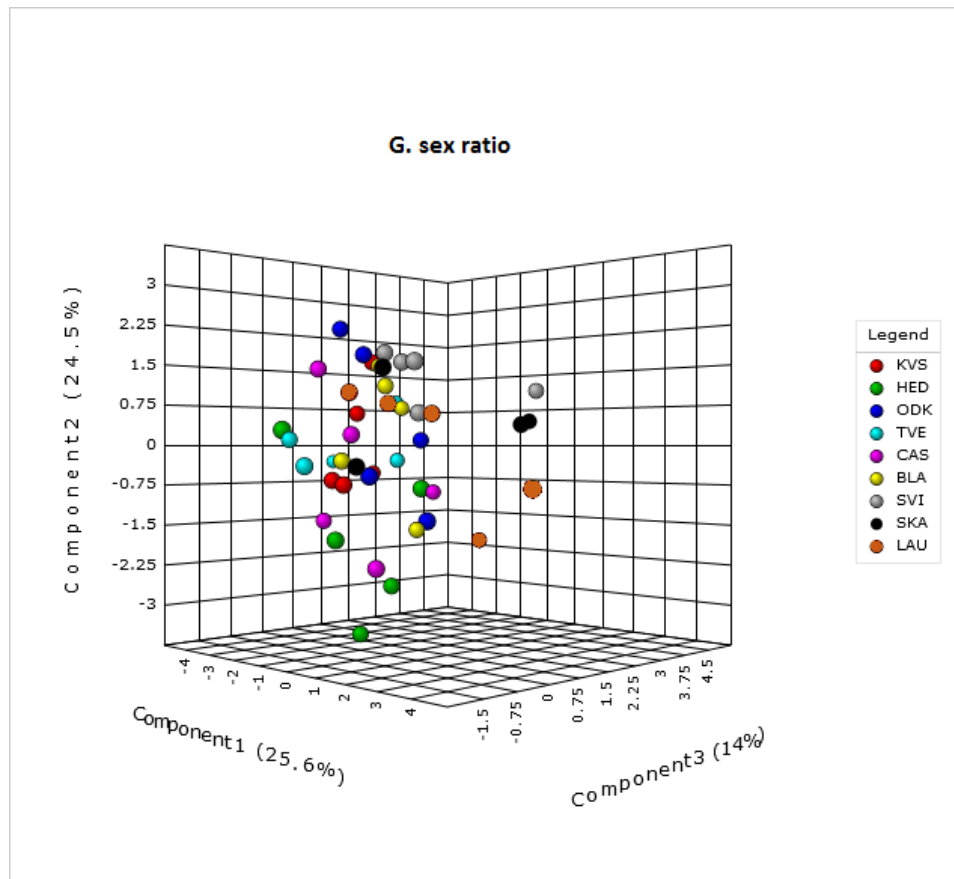
598 **A.5 overview of the VIP scores resulting from the PLS-DA**

599 An overview off all metabolites with a VIP score of at least 1.2 (1.0 for egg survival because removing the metabolites with a score

600 between 1.0 and 1.2 resulted in a decreased percentage of variation explained) for component 1 and/ or 2 in the PLS-DA performed for

601 latitude (A) and each of six life-history traits (daily fecundity (B), lifetime fecundity (C), egg survival (D), longevity (E), dispersal

602 propensity (F) and sex ratio(G)). Note that in the tables below, the VIP scores for the original dataset (containing all forty-three

603 identified metabolites) are given. The order of the metabolites (from high to low VIP scores) in these tables can therefore deviate from

604 the order in Fig. 3 (main text), as Fig. 3 is based on VIP scores of the final dataset (after removal of all non-explanatory metabolites).

605 **A. latitude**

metabolite	Comp. 1	Comp. 2
Proline	1.6583	1.0501
Tryptophan	1.5937	1.0289
Inositol	1.5831	1.0274
Malic acid	1.4947	0.9356
Lysine	1.4669	0.9290
Valine	1.4577	0.9261
Isoleucine	1.4219	0.9054
Methionine	1.4115	0.8879
Glycine	1.4090	0.8936
Aspartic acid	1.3881	0.8733
Threonine	1.3705	0.9118
Alanine	1.3572	0.8500
Leucine	1.3414	0.9084
Quinic acid	1.3313	1.6940
Glucose-6-phosphate	0.6854	2.0669
Sorbitol	1.0691	1.9186
Fructose-6-phosphate	0.4123	1.7967
Ribose	0.8471	1.2669

606

607 **B. daily fecundity**

metabolite	Comp. 1	Comp. 2
Glycerol	1.7613	1.8109

Phenylalanine	1.6599	1.2292
Sorbitol	1.6526	1.4537
Valine	1.5163	1.1636
Xylitol	1.4767	1.2651
Lysine	1.4179	1.0710
Isoleucine	1.4127	1.1010
Glutamic acid	1.3973	0.9667
Tyrosine	1.3441	1.0950
Leucine	1.3261	1.1045
Threonine	1.2862	1.0977
Inositol	1.2444	0.9216
Glycine	1.2215	0.9496
Tryptophan	1.2095	0.9172

608

609 **C. lifetime fecundity**

metabolite	Comp. 1	Comp. 2
Lysine	1.5753	0.8880
Glycine	1.5660	0.8728
Tryptophan	1.5312	0.8706
Inositol	1.5101	0.8566
Valine	1.4669	0.9187
Proline	1.4248	0.8628
Alanine	1.4119	0.7920
Isoleucine	1.3774	0.8873
Malic acid	1.3485	0.8655
Phenylalanine	1.3188	0.9131
Leucine	1.3161	0.9319
Lactic acid	1.2638	1.3055
Glucose-6-phosphate	1.2298	1.7909
Quinic acid	1.2139	0.9298
Fructose-6-phosphate	0.9630	1.6380

610

611 **D. egg survival**

metabolite	Comp. 1	Comp. 2
Glutamic acid	1.7874	1.3641
Alanine	1.6795	1.1028
Tyrosine	1.4029	0.9337
Phenylalanine	1.3828	0.9234
Fructose-6-phosphate	1.3581	0.8561
Glucose-6-phosphate	1.3492	0.8500
Methionine	1.2596	0.8576
Xylitol	1.2585	1.0527
Glycerol-3-phosphate	1.2193	0.7804
Citric acid	1.2164	0.8394
Serine	1.1977	0.9293
GABA	1.1951	2.2339
Gluconolactone	1.1857	0.8810
Adonitol	1.1823	0.8112
Succinic acid	1.1628	1.0936
Leucine	1.1035	0.9157
Threonine	1.0640	0.9255
Valine	1.0216	0.8728
Ornithine	1.0141	0.6404
Ethanolamine	0.8772	1.9297
Glycerol	0.5981	1.7154
Fumaric acid	0.0152	1.5159
Glutamic acid	1.7874	1.3641
Glyceric acid	2.1181	1.3224

612

613   **E. longevity**

metabolite	Comp. 1	Comp. 2
Glucose	2.1463	1.4658
Succinic acid	1.9802	1.3826
Sorbitol	1.9329	1.3478
Citric acid	1.6265	1.5863
Glycine	1.4818	1.1018

47

Galactose	1.4447	1.1659
Lysine	1.4091	1.0542
Lactic acid	1.3965	1.0108
Fructose-6-phosphate	1.3276	1.4043
Glucose-6-phosphate	1.3168	1.3696
Phosphoric acid	1.2220	1.0486

614

615 **F. dispersal propensity**

metabolite	Comp. 1	Comp. 2
Putrescine	1.9187	1.2144
Phosphoric acid	1.7678	1.0416
Lysine	1.7503	1.0280
Methionine	1.7343	1.0303
Xylitol	1.7230	1.9394
Phenylalanine	1.5574	0.9821
Fructose	1.4964	0.9235
Tryptophan	1.2684	0.7950
Glycerol-3-phosphate	1.2602	0.7514
Leucine	1.2553	0.9363
Inositol	1.2549	0.7824
Glucose-6-phosphate	0.6820	1.4379
Mannose	0.2110	1.3947

616

617 **G. sex ratio**

metabolite	Comp. 1	Comp. 2
GABA	2.7316	2.5800
Succinic acid	2.5829	2.4256
Glutamic acid	2.0087	1.8893
Glucose-6-phosphate	1.7334	1.6527
Sorbitol	1.5872	1.4965
Citric acid	1.4820	1.4001

Fructose-6-phosphate	1.4541	1.3954
Xylitol	1.2478	1.3554
Ethanolamine	1.2072	1.3270
Fructose	0.9418	0.9842*

618   *After removal of the least explanatory metabolites in the dataset, the VIP score of fructose changed to >1.2. Fructose was therefore

619   retained for further analysis.

620

621   ## A.6 overview of all linear regressions

622   An overview of all linear regressions according to latitude (A) and each of the six life-history traits (daily fecundity (B), lifetime fecundity

623   (C), egg survival (D), longevity (E), dispersal propensity (F) and sex ratio(G)) known to covary with latitude in the study species (see Van

624   Petegem *et al.* 2016). The direction of change (correlation) is each time given, together with the F- and *p*-value of the linear regression.

625   The degrees of freedom used in determining the F-value (Num DF and Den DF) are also shown.

626

627   ### A. latitude

metabolite	correlation	Num DF	Den DF	F	p
Proline	negative	1	7	9.85	**0.0164**
Tryptophan	negative	1	7	6.48	**0.0383**
Inositol	negative	1	7	6.43	**0.0389**
Aspartic acid	negative	1	7	6.62	**0.0369**
Glycine	negative	1	7	7.19	**0.0315**
Malic acid	negative	1	7	6.59	**0.0372**
Valine	negative	1	7	8.15	**0.0245**
Lysine	negative	1	7	3.77	0.0932
Alanine	negative	1	7	4.16	0.0808
Methionine	negative	1	7	3.06	0.1238
Isoleucine	negative	1	7	8.81	**0.0209**
Quinic acid	positive	1	7	5.59	**0.0500**
Threonine	negative	1	7	7.35	**0.0301**
Leucine	negative	1	7	6.71	**0.0359**
Sorbitol	positive	1	7	1.13	0.3221
Glucose-6-phosphate	positive	1	7	1.00	0.3511
Fructose-6-phosphate	positive	1	7	0.34	0.5789

628

629   ### B. daily fecundity

metabolite	correlation	Num DF	Den DF	F	p
Glycerol	negative	1	7	1.91	0.2096
Phenylalanine	positive	1	7	1.87	0.2135
Glutamic acid	positive	1	7	0.47	0.5193
Sorbitol	negative	1	7	3.37	0.0947
Valine	positive	1	7	2.41	0.1648
Isoleucine	positive	1	7	2.46	0.1606
Lysine	positive	1	7	1.34	0.2588
Tyrosine	positive	1	7	1.31	0.2897
Leucine	positive	1	7	1.54	0.2553
Threonine	positive	1	7	1.83	0.2182
Xylitol	negative	1	7	0.46	0.5193
Glycine	positive	1	7	1.51	0.2587
Inositol	positive	1	7	0.71	0.4284
Tryptophan	positive	1	7	0.62	0.4584

630

631   **C. lifetime fecundity**

metabolite	correlation	Num DF	Den DF	F	p
Lysine	positive	1	7	2.57	0.1531
Glycine	positive	1	7	3.86	0.0903
Tryptophan	positive	1	7	0.60	0.4629
Inositol	positive	1	7	0.61	0.4592
Valine	positive	1	7	1.61	0.2455
Proline	positive	1	7	0.36	0.5665
Alanine	positive	1	7	0.15	0.7089

Isoleucine	positive	1	7	1.47	0.2653
Malic acid	positive	1	7	1.24	0.3026
Phenylalanine	positive	1	7	0.70	0.4317
Leucine	positive	1	7	0.91	0.3717
Lactatic acid	positive	1	7	0.22	0.6503
Glucose-6-phosphate	negative	1	7	2.38	0.1671
Quinic acid	negative	1	7	0.58	0.4716
Fructose-6-phosphate	negative	1	7	2.15	0.1862
Glycerol-3-phosphate	negative	1	7	1.07	0.3363

632

633 **D. egg survival**

metabolite	correlation	Num DF	Den DF	F	p
Glutamic acid	negative	1	6	3.79	0.0994
Alanine	negative	1	6	4.40	0.0808
Tyrosine	negative	1	6	5.24	0.0621
Phenylalanine	negative	1	6	2.70	0.1516
Fructose-6-phosphate	negative	1	6	4.55	0.0770
Glucose-6-phosphate	negative	1	6	4.66	0.0742
Methionine	negative	1	6	1.84	0.2238
Xylitol	negative	1	6	0.37	0.5649
Succinic acid	negative	1	6	3.02	0.1327
Glycerol-3-phosphate	negative	1	6	4.70	0.0733
Citric acid	negative	1	6	4.61	0.0754
Serine	negative	1	6	5.03	0.0661
GABA	positive	1	6	0.38	0.5580

Adonitol	negative	1	6	9.07	**0.0237**
Gluconolactone	negative	1	6	0.60	0.4697
Leucine	negative	1	6	3.16	0.1260
Threonine	negative	1	6	2.85	0.1425
Valine	negative	1	6	2.23	0.1860
Ornithine	negative	1	6	2.68	0.1525
Ethanolamine	negative	1	6	0.00	0.9891
Glycerol	positive	1	6	0.03	0.8651
Malic acid	negative	1	6	2.22	0.1871
Glyceric acid	negative	1	6	0.94	0.3698
Fumaric acid	negative	1	6	0.44	0.5305

634

635 **E. longevity**

metabolite	correlation	Num DF	Den DF	F	p
Glucose	positive	1	7	0.52	0.4955
Succinic acid	negative	1	7	0.00	0.9482
Sorbitol	positive	1	7	0.05	0.8306
Citric acid	negative	1	7	2.14	0.1870
Glycine	positive	1	7	2.83	0.1365
Galactose	positive	1	7	1.13	0.3236
Lysine	positive	1	7	1.65	0.2404
Lactic acid	positive	1	7	0.89	0.3777
Fructose-6-phosphate	negative	1	7	1.16	0.3174
Glucose-6-phosphate	negative	1	7	1.16	0.3172
Phosphoric acid	positive	1	7	0.97	0.3573

Ornithine	positive	1	7	0.79	0.4023
Ribose	positive	1	7	0.23	0.6430
Xylitol	negative	1	7	0.07	0.7935
Ethanolamine	positive	1	7	1.09	0.3310
Glycerol	positive	1	7	0.06	0.8184
Glutamic acid	positive	1	7	0.18	0.6826
Mannose	positive	1	7	0.88	0.3804
Glycerol-3-phosphate	negative	1	7	0.61	0.4607
Methionine	positive	1	7	0.05	0.8332
Arabitol	positive	1	7	0.15	0.7066
Serine	positive	1	7	0.15	0.7117
Adonitol	positive	1	7	0.00	0.9879

636

637 **F. dispersal propensity**

metabolite	correlation	Num DF	Den DF	F	p
Putrescine	negative	1	7	8.64	**0.0218**
Phosphoric acid	negative	1	7	7.33	**0.0303**
Lysine	negative	1	7	6.38	**0.0394**
Methionine	negative	1	7	8.16	**0.0245**
Xylitol	positive	1	7	0.66	0.4436
Phenylalanine	negative	1	7	9.05	**0.0197**
Fructose	negative	1	7	7.29	**0.0307**
Tryptophan	negative	1	7	5.10	0.0585
Glycerol-3-phosphate	negative	1	7	3.67	0.0969
Leucine	negative	1	7	8.14	**0.0246**

Inositol	negative	1	7	4.94	0.0615
Glucose-6-phosphate	positive	1	7	0.37	0.5608
Mannose	negative	1	7	0.02	0.9013

638

639 **G. sex ratio**

| metabolite	correlation	Num DF	Den DF	F	p
GABA	negative	1	7	3.26	0.1141
Succinic acid	positive	1	7	3.49	0.1038
Glutamic acid	positive	1	7	0.91	0.3716
Glucose-6-phosphate	positive	1	7	2.60	0.1511
Sorbitol	positive	1	7	0.83	0.3932
Citric acid	positive	1	7	1.82	0.2193
Fructose-6-phosphate	positive	1	7	1.91	0.2089
Xylitol	negative	1	7	0.26	0.6281
Ethanolamine	negative	1	7	1.20	0.3091
Fructose	negative	1	7	0.70	0.4294

640

641

642 **A.7 pathway maps**

643 Pathway maps of the two pathways in which some of the metabolites identified in the GC-MS analysis (visualised with either red or blue

644 rectangles) play an important role. In the aminoacyl-tRNA biosynthesis pathway, the metabolites in the red rectangles correlate

645 negatively with latitude and the blue ones correlate negatively with dispersal propensity. In the valine, leucine and isoleucine

646 biosynthesis pathway, the metabolites in the red rectangles correlate negatively with latitude. The maps were used from KEGG (Kyoto

647 Encyclopedia of Genes and Genomes) (Kanehisa *et al.* 2015).

AMINOACYL-tRNA BIOSYNTHESIS

VALINE, LEUCINE AND ISOLEUCINE BIOSYNTHESIS

00290 9/13/12
(c) Kanehisa Laboratories

649

58
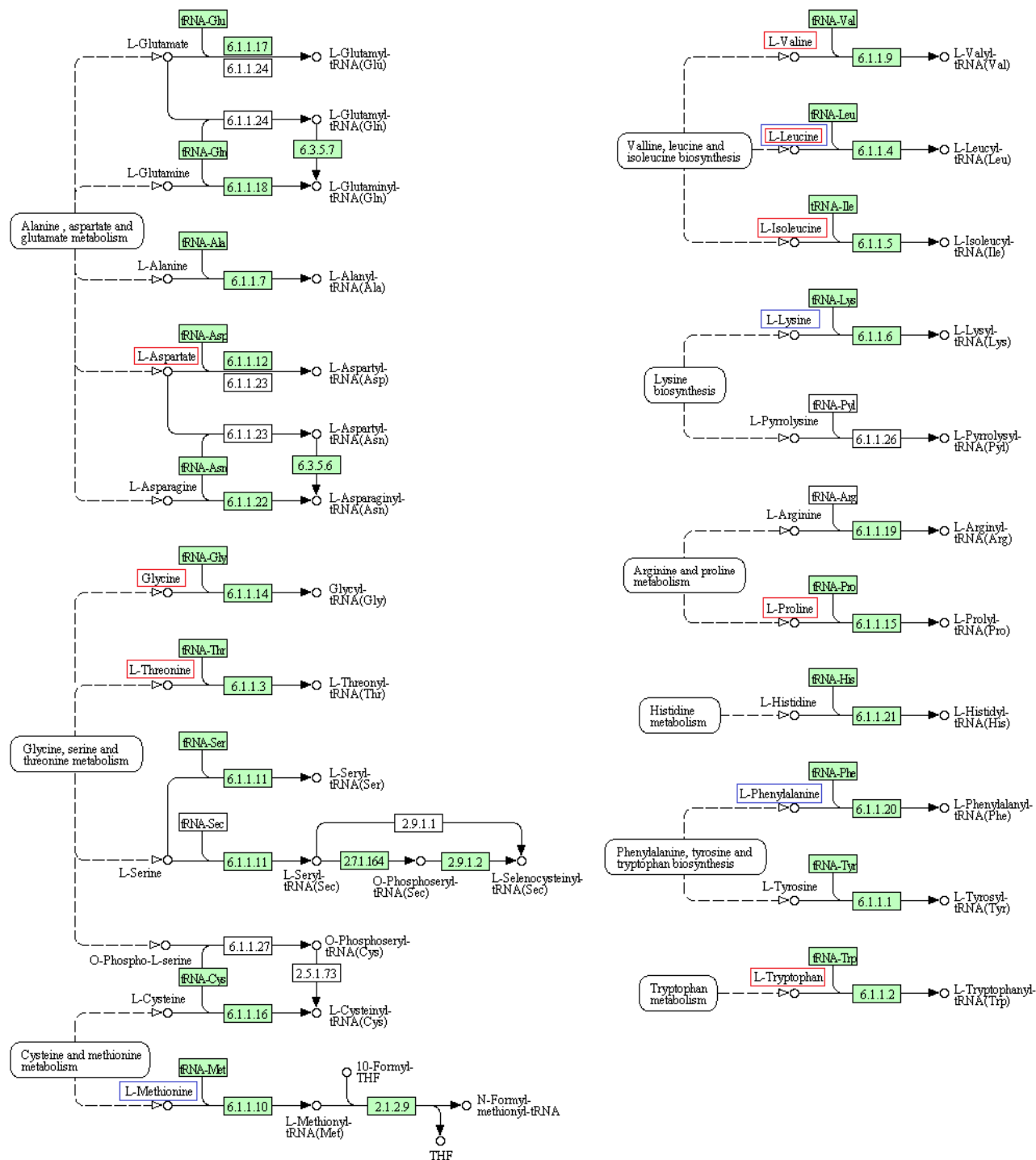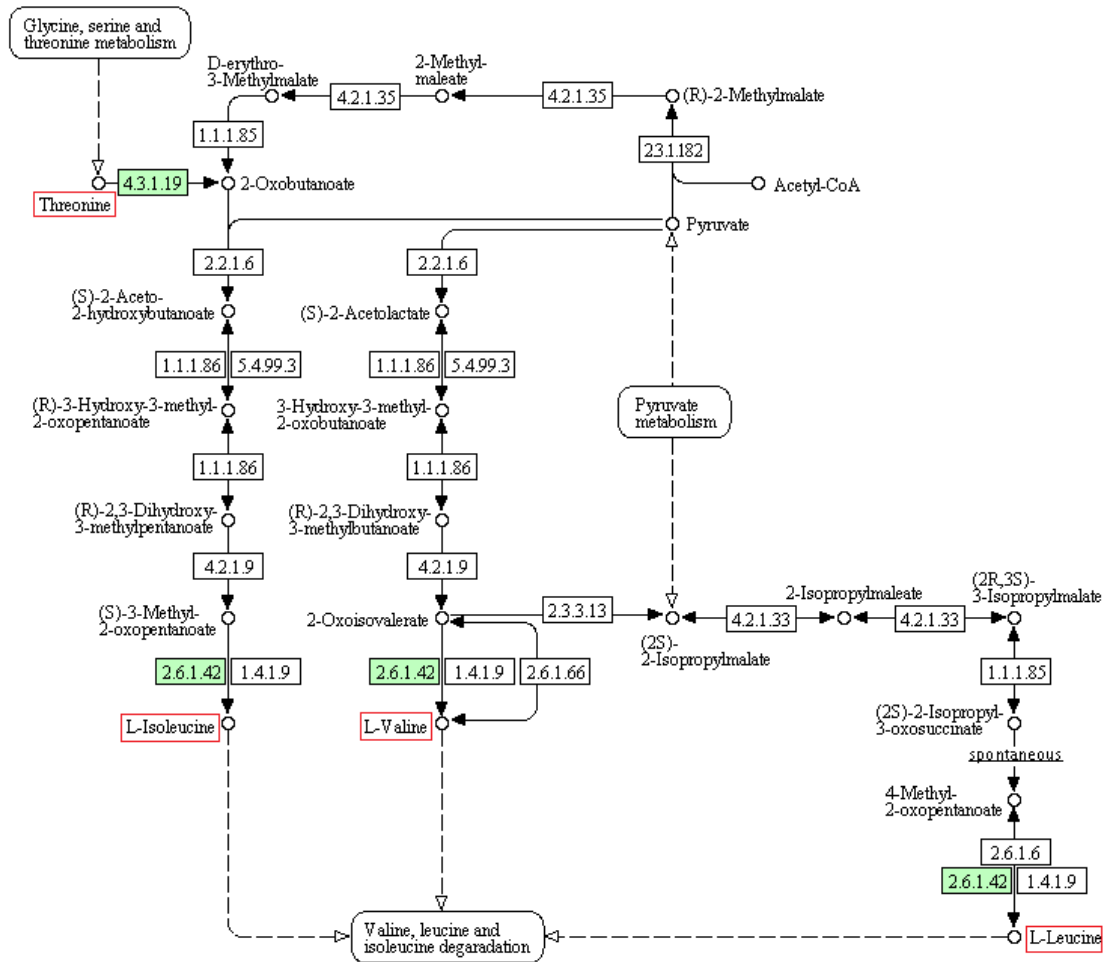
650   **A.8 regression for all populations *versus* for *L. periclymenum* only**

651   For those metabolites that gave significant results (see appendix A.6) for latitude or a specific life-history trait (daily fecundity, lifetime

652   fecundity, egg survival, longevity, dispersal propensity or sex ratio) in the regressions run for all nine populations (hence for *L.*

653   *periclymenum*, *Euonymus europaeus*, *Sambucus nigra* and *Humulus lupulus*), we repeated these regressions but then for *L.*

654   *periclymenum* only (hence for only six populations –see appendix A.1). The table gives the slopes and p-values of both the regressions

655   run for all populations (slope_all and p_all) and those run for *L. periclymenum* only (slope_*per* and p_*per*).

656

trait	metabolite	slope_*per*	slope_all	p_*per*	p_all
latitude	Proline	-0.04354	-0.02733	0.0205	0.0164
latitude	Tryptophan	-0.05170	-0.04474	0.1766	0.0383
latitude	Inositol	-0.05628	-0.04890	0.1880	0.0389
latitude	Aspartic acid	-0.01956	-0.02840	0.2757	0.0369
latitude	Glycine	-0.00681	-0.03254	0.6500	0.0315
latitude	Malic acid	-0.00703	-0.03858	0.6992	0.0372
latitude	Valine	-0.02914	-0.02727	0.1726	0.0245
latitude	Isoleucine	-0.03947	-0.02997	0.0947	0.0209
latitude	Quinic acid	0.02094	0.01594	0.2119	0.0500
latitude	Threonine	-0.02807	-0.02700	0.2046	0.0301
latitude	Leucine	-0.01727	-0.02270	0.3466	0.0359
egg survival	Adonitol	-0.00651	-0.00258	0.0751	0.0237
dispersal propensity	Putrescine	-0.00092	-0.00108	0.0854	0.0218
dispersal propensity	Phosphoric acid	-0.10219	-0.13336	0.0492	0.0303
dispersal propensity	Lysine	-0.00277	-0.00553	0.0942	0.0394
dispersal propensity	Methionine	-0.00793	-0.01081	0.1461	0.0245
dispersal propensity	Phenylalanine	-0.01652	-0.02201	0.1208	0.0197
dispersal propensity	Fructose	-0.00266	-0.00221	0.0744	0.0307
dispersal propensity	Leucine	-0.03211	-0.04462	0.1224	0.0246

657

59

658   **References**

659

660   Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. (2015) KEGG as a

661   reference resource for gene and protein annotation. *Nucleic Acids Research*.

662   Van Petegem, K., Boeye, J., Stoks, R. & Bonte, D. (2016) Spatial selection and local adaptation

663   jointly shape life-history evolution during range expansion. *bioRxiv*.

664   Waller, M., Whitney, R., Taylor, M., Miyagawa, H., Matsuda, K. & Nakagawa, K. (2007)

665   Multiresidue pesticides analysis using synchronous Scan/SIM mode GC-MS. *Lc Gc*

666   *North America*, 39-39.

667

668